

1-1-1995

Industrial computer vision using undefined feature extraction

Phil Evans

University of Wollongong

John A. Fulcher

University of Wollongong, john@uow.edu.au

Philip Ogunbona

University of Wollongong, philipo@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>



Part of the [Physical Sciences and Mathematics Commons](#)

Recommended Citation

Evans, Phil; Fulcher, John A.; and Ogunbona, Philip: Industrial computer vision using undefined feature extraction 1995, 1145-1149.

<https://ro.uow.edu.au/infopapers/2153>

Industrial computer vision using undefined feature extraction

Abstract

This paper presents an application of computer vision. The implementation and operation of the system is vision in a real-world uncontrolled environment found at BHP Steel Port Kembla. The task is visual identification of torpedo ladles at a Blast Furnace. The task is achieved by reading numbers attached to each ladle. IMPLEMENTATION Number recognition is achieved through use of feature extraction using a Multi-Layer Perceptron (MLP) Artificial Neural Network (ANN). The novelty in the method used in this application is that the features the MLP is being trained to extract are undefined before the MLP is initialised. The results of the MLP processing are passed to a decision tree for analysis and final classification of each object within the image. This technique is achieving a recognition rate on previously unseen images of greater than 80%.

Keywords

feature, undefined, extraction, vision, industrial, computer

Disciplines

Physical Sciences and Mathematics

Publication Details

Evans, P., Fulcher, J. A. & Ogunbona, P. (1995). Industrial computer vision using undefined feature extraction. IEEE International Conference on Neural Networks - Conference Proceedings: Vol 2 (pp. 1145-1149).

INDUSTRIAL COMPUTER VISION USING UNDEFINED FEATURE EXTRACTION

Phillip Evans
BHP Information Technology Pty Ltd.
evans.phil.pm@bhp.com.au

John Fulcher
Department of Computer Science
University of Wollongong
john@cs.uow.edu.au

Phillip Ogunbona
Department of Electrical and Computer Engineering
University of Wollongong
pogna@elec.uow.edu.au

1. ABSTRACT

This paper presents an application of computer vision in a real-world uncontrolled environment found at BHP Steel Port Kembla. The task is visual identification of torpedo ladles at a Blast Furnace which is achieved by reading numbers attached to each ladle.

Number recognition is achieved through use of feature extraction using a Multi-Layer Perceptron (MLP) Artificial Neural Network (ANN). The novelty in the method used in this application is that the features the MLP is being trained to extract are undefined before the MLP is initialised.

The results of the MLP processing are passed to a decision tree for analysis and final classification of each object within the image.

This technique is achieving a recognition rate on previously unseen images of greater than 80%

2. INTRODUCTION

Current computer vision techniques [1, 2] tend to require a very controlled operating environment in order to maintain reliable operation. The aim of this work is to develop a solution to a significant number of industrial vision problems in an uncontrolled environment. Having achieved that, a further goal is to make the installation and operation of the equipment as simple as possible, so as to facilitate installation and operation by a professional person who is not necessarily an expert in computer vision.

This paper outlines the procedure and some early results in the implementation towards the second goal. The implementation and operation of the system is discussed.

3. IMPLEMENTATION

In order for the vision system to automatically detect and identify objects within its field of view, it must undergo several stages of preprocessing, and use some form of artificial intelligence to choose and store the "interesting" objects. At this stage the system operates only in two dimensions, but there is potential for it to be expanded into three dimensions.

There are three basic operating stages (Figure 1):

- Preprocessing (using conventional mathematical techniques)
- Feature Construction (Using Artificial Neural Networks)
- Object Construction and Identification (Using Machine Learning Techniques)

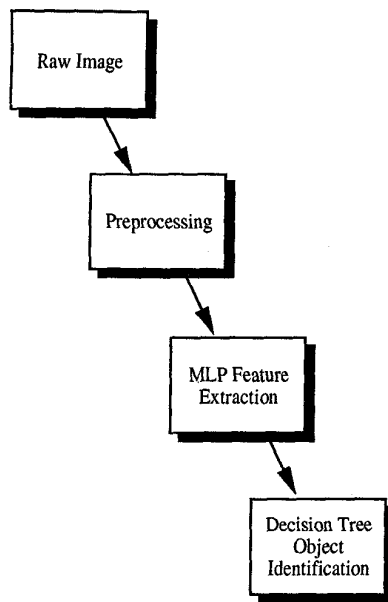


Figure 1. System Block Diagram

3.1 Preprocessing

Each image is preprocessed by the relative brightness of each colour for each pixel (Figure 2). For a given pixel colour (R_1, G_1, B_1) all pixels with a colour within a bounding region are extracted and used in turn to extract regions [3, 4].

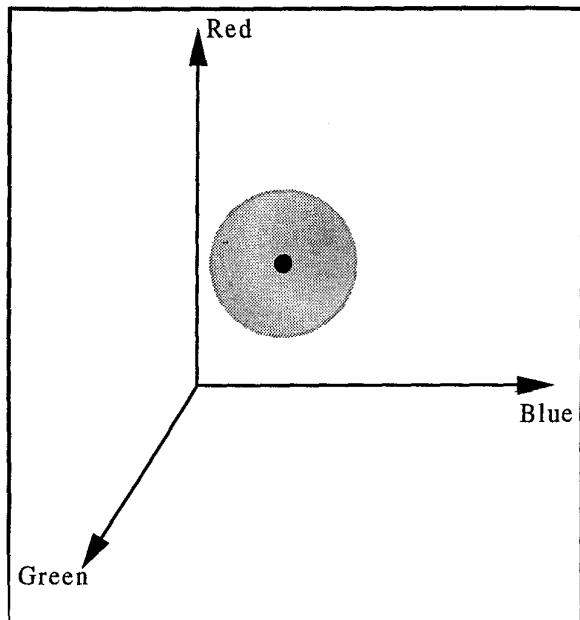


Figure 2. Like Colour Extraction

This process is extended to all pixels in the image to extract all of the similarly coloured objects in the image. Images which are too small to have sufficient pixel resolution for recognition purposes are discarded. Regions of similar colour are extracted and then stored as objects.

These objects are further filtered based upon information already known about the objects of interest. The filtering criteria currently in use includes the ratio of object height to width, and approximate colour.

3.2 Feature Extraction

Feature Extraction is performed by a set of MLP Artificial Neural Networks (ANN). Each individual region is examined for dimension and orientation before being scaled and fed into the appropriate MLP. The result from this stage is the set of (undefined) features present in each region in the image.

3.3 Object Construction and Identification

With the features defined for each region in the image, object construction can now take place. To allow for more complex objects composed of more than one colour (or shade), each region is examined with respect to every other region to identify any adjacent edges and thus establish possible inter-region relationships. The system does not as yet fully deal with occlusion, although that does not mean that the MLPs can not allow for some degree of occlusion.

A simple machine learning technique is now utilised to recognise the objects within the image. Each MLP classification is fed into the decision tree, with the associated region classifications also providing input to the decision process. An operator specified number of sample images for each type of classification is stored for later classification by the operator.

The operator later examines the sample images and verifies that the system has identified as separate classifications the objects of interest. The operator also manually prunes the Decision Tree through identifying similar classifications in different branches of the decision tree. Uninteresting classifications can also be removed so that only the truly interesting objects are identified during run-time operation of the system.

4. TEST ENVIRONMENT

4.1 BHP Steel #5 Blast Furnace, Port Kembla

BHP Steel, Slab and Plate Products Division Port Kembla have a long standing system for identifying torpedo ladles at the #5 Blast Furnace. At the Blast Furnace, solid material is melted down into molten iron. This molten iron is then "tapped", or poured, into a torpedo ladle which is, in essence, a train carriage. The torpedo ladle is used to transport the molten material to the next stage in the steelmaking process, the Basic Oxygen Steelmaking (BOS) plant. The blast furnace consists of one vessel with three (3) casting floors in operation. Each casting floor supports two (2) bays where the torpedo ladles are placed to be charged. The bays are paired, so that while one torpedo ladle is being charged, the torpedo ladle in the second bay can be replaced. In this way a continuous tapping operation can proceed.

Each torpedo ladle is identified by a two digit identifier found on either end, and on the sides, of the ladle. The numbers are cut out of metal plate, and stand about 300mm tall. The numbers are repainted white, and are well maintained, so that the ladles can be easily identified by the train drivers.

The proposal was to replace the current tracking system with a computer vision system based upon the techniques described earlier.

An earlier tracking system utilised Radio Frequency (RF) tags on each end of each ladle. When the tag is polled by an RF reader, it responds with a unique identifier which can then be associated with a particular ladle identifier. This is the system by which the torpedo ladles are automatically tracked as they enter and exit the bays under the blast furnace. The problem with using this type of technology is that the equipment has to be placed in a very unfriendly environment. When there is a spillage on a ladle, the RF tag can be melted. In cases where there is an extreme spillage and molten material ends up on the floor, it is not uncommon for the RF readers to be damaged, and require replacement.

Computer vision technology offers a non-contact solution where the sensing devices, and the ladle identifiers, can be located in a safe location such that they will not be affected by any spillage.

4.2 The Operating Environment

The charge bays are in operation 24 hours a day and only partially enclosed. They are exposed to the outside environment and are subject to changing light conditions, dust, and steam clouds blown by the wind.

The numbers themselves are well maintained, being repainted white approximately every 3 months. However, the numbers are made from steel plate and are sometimes pushed over, or into each other, or mangled in some other manner.

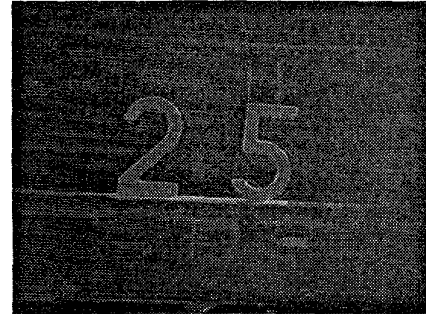


Figure 3. Example Night time Image

5. METHOD

For each new application, the ANN and Decision Tree Construction stages need to be customised. The MLP needs to be configured and trained first, with the Decision Tree being used to classify the output from the network.

5.1 MLP Training

In this instance, the MLP was configured with an input layer of 432 ($=18 \times 24$) nodes, second layer of 150 nodes, third layer of 75 nodes, and an output layer of 20 nodes. The weights were initialised with random values ranging between -0.1 and 0.1.

The MLP is trained in a technique where the training data is generated from the unclassified input data. To begin, a NULL vector is presented to the ANN, and the result stored. Then each training vector is presented to the untrained ANN, and the resultant output compared to the output generated by the NULL vector. Individual output points which are higher in magnitude than the output produced by the NULL vector are set to 1, and the output points lower in magnitude are set to 0. This defines the set of "features" as identified by the untrained ANN which define each input image. At this point a "sanity check" is performed to ensure that the ANN has, in fact, identified different features for each classification (ie: not all training images were classified in the same way). So far, this technique has not yielded such a problem, probably due to the large size of the MLP. If the generated training data is OK, training is commenced.

The training curve produced when training an MLP in this fashion is different to that normally encountered during training. An example is shown below. This is believed to be because the MLP is (effectively) initialised with correct weights, and the information it already contains is simply further enhanced. When using the more conventional approach, the MLP is not initialised with useful information and thus must endure a period of “unlearning” before it can learn the required responses.

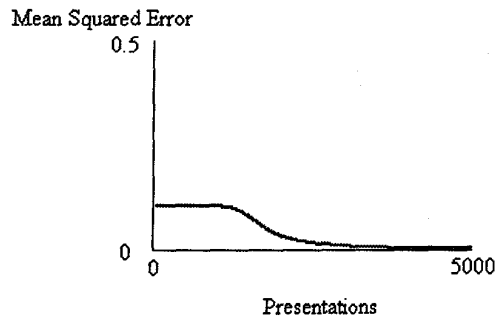


Figure 4. Typical MLP training error curve

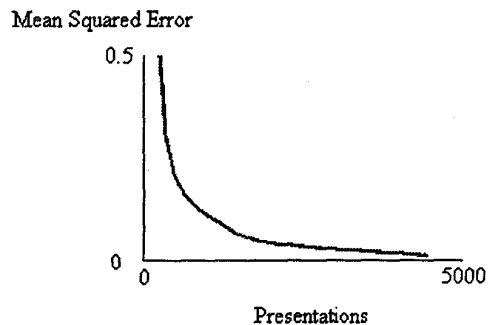


Figure 5. Baseline MLP training error curve

In order to improve the representation in the ANN, it is possible to re-iterate the training procedure on the now partially-trained MLP in order to better define the features that it classifies, and how they relate to the input data set. This helps to reduce the variation in the target classifications and decreases the training time. By re-presenting the original training data to the partially trained MLP for re-classification, it is possible to improve the reliability of the MLP's feature extraction.

5.2 Decision Tree Creation

When training is complete, the original unclassified data is again passed through the MLP to gain the final output classifications for each data type. At the same time, a Decision Tree is generated to classify the outputs according to the features present.

Once the Decision Tree is built, the supervisor must manually classify each resultant branch of the tree. The tree is then manually pruned so that each classification uses as many common features as possible. This will both reduce the size of the tree, and enable better generalisation on the MLP output. It should be noted that this does not mean that there will be only one branch in the tree for each

6. RESULTS

In order to test this technique, the system was trained on a set of 10 digits (0..9) typed in 10 different fonts. Some example training images are shown in Figures 6 and 7.

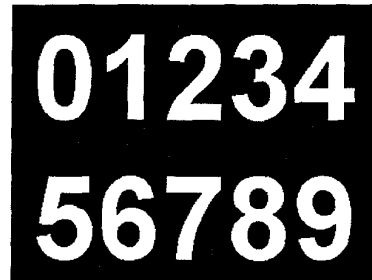


Figure 6. MS Sans Serif Font

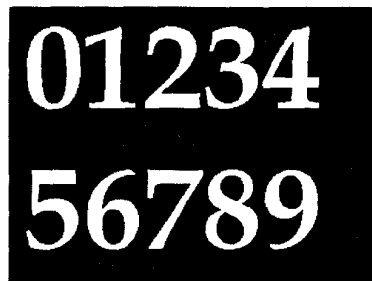


Figure 7. Courier New Font

Using the standard back-propagation algorithm, the MLP was trained and the decision tree generated. During the first training iteration, 5000 exemplars were presented, and the error was reduced to 0.2%. The unclassified training exemplars were re-presented, and the MLP was trained for a further 5000 exemplars to reduce the error below 0.05%.

The decision tree generated at the end of the training period was twelve (12) levels deep, but this was pruned back to a depth of five (5) levels.

When tested with the training data, the system achieved 100% recognition accuracy. This dropped to between 80% and 90% when presented with previously unseen data (ie: images not previously met).

6.1 Test Results

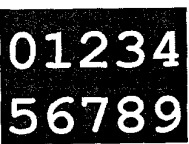
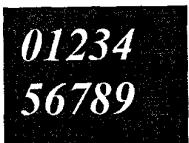
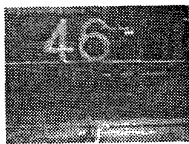
Typeface	Sample Image	Recognition Accuracy
Plain text (Courier New Font)		100%
Italic typeface (Times New Roman)		80%
Real images		88%

Figure 8. Results

7. FURTHER WORK

The results achieved to date suggest that further refining of the MLP training procedure would result in better overall performance. The inclusion of italicised text in the training set, as well as some sample images from the real-world application at the Blast Furnace should improve the recognition accuracy on these types of characters.

If the recognition accuracy of this system can be increased to above 95% in all cases, the system will be implemented into the production environment at #5 Blast Furnace.

The next stage of work will extrapolate the MLP feature extraction ability to allow the system to identify objects previously unseen by the MLP, without the need for retraining. The new image will be presented to the system, and the results from the MLP used to generate new branches within the decision tree. The tree could then be further refined to include knowledge about which features define the new objects, and theoretically be able to identify them.

8. CONCLUSIONS

The results to date validate the feasibility of this technique for identifying objects within images. The MLP is capable of extracting self-defined features and using these features to identify individual objects. Potentially this technique will be able to be extrapolated to a system which can identify the feature set which define an object and identify similar objects without the need for any further training.

9. REFERENCES

- [1] Anil K. Jain, "Fundamentals of Digital Image Processing", Prentice-Hall, Inc. 1989.
- [2] William K. Pratt, "Digital Image Processing - Second Edition", Psychophysical Vision Properties, John Wiley & Sons, Inc. 1991.
- [3] G. Stewart, D. Fraser, "FaceART - ART2 in Face Recognition Using Raw Pixel Data", Proceedings of the Fifth Australian Conference on Neural Networks, Brisbane, January 1994.
- [4] T. Dunstone., "Automatic Head Extraction", Proceedings of the Fifth Australian Conference on Neural Networks, Brisbane, January 1994.